Change-Point Detection in Moving Average Model Using Reversible Jump MCMC Algorithm

By Suparman

International Journal of Engineering & Technology, 5 (x) (2017) xxx-xxx

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi:
Research paper, Short communication, Review, Technical paper

Change-Point Detection in Moving Average Model Using Reversible Jump MCMC Algorithm

Suparman[1]*

[1]*Department of Mathematics Education, University of Ahmad Dahlan, Indonesia*
**Corresponding author E-mail: suparman@pmat.uad.ac.id*

Abstract

This paper aims to estimate a parameter of piecewise MA model that has a gaussian white noise. A Bayesian method is adopted. A prior distribution of the parameter of piecewise MA model is selected and then this prior distribution is combined with a likelihood function of data to get a posterior distribution. Based on this posterior distribution, a Bayesian estimator for the parameter of piecewise MA model is estimated. Because the order of MA model is considered a parameter, a method of reversible jump Markov Chain Monte carlo (MCMC) is adopted. The result is an estimation of parameter MA model that can be simultaneously calculated.

Keywords: *Bayesian, Change-point, Moving Average, Reversible Jump MCMC.*

1. Introduction

A moving average (MA) model is a mathematical model that is widely used in many fields. Nakamura [1] used the MA model to estimate the natural location of the bee colony. Wang [2] used the MA model as an error in this non-linear-finite impulse response (IIR). Various authors have examined the estimation of MA model parameters, e.g. Dimitriou [3] proposed the Yule-Walker estimator to estimate the parameters of the MA model. But in many applications, many data have models that change from one time interval to another. Combined several mathematical models are needed to model data that has models that change from one time interval to another time interval.
The piecewise model is a mathematical model developed to model the data that changes patterns from one time interval to another time interval. The piecewise model is widely used in many fields. Hong [4] used piecewise regression model on semiconductors. Diallo [5] used the piecewise model to model longitudinal data. Tang [6] and Tang [7] used the piecewise model to model the virus-immune system. Chamrouki [8] used the piecewise regression model for clustering and segmentation. Xu [9] used the piecewise model for the diversity of species in the river. Buscot [10] used the piecewise regression model to detect the divergence of trajectory in groups of children with different developmental phases. Hector [11] used the piecewise model to calculate energy expenditure (EE) in real time from heart rate. Soleimanmeigouni [12] used a linear piecewise model to model the degradation of railway geometry with known break points. But in various applications, the number of models in the piecewise model is generally unknown. The number of models is a parameter that needs to be estimated based on the data.
The MCMC reversible jump algorithm [13] is a method that can be used for the selection of mathematical models in which the number of mathematical models is assumed to be unknown. Suparman [14] used reversible jump MCMC algorithm to estimate ARMA model parameters of unknown model order. Suparman [15] implemented the MCMC reversible jump algorithm to detect piecewise regression models where the number of change-points is unknown.
Moving-Average (MA) model is a model that is often used in many fields. If the piecewise MA model is matched against the signal data, the model parameters will be generally unknown. The objective of this paper is to estimate the parameters of the piecewise MA model. There are so many piecewise MA models. In this paper, the noise distribution for each segment will be assumed as the Gaussian distribution with mean 0 and unknown variance.
Consider a signal $y = (y_1,\dots,y_n)$ where n is the number of observations, modelled as an MA model with piecewise constant parameter and k ($k = 0,\dots,k_{max}$) change-points. Mathematically, the model is the following:

$$y_t = z_t + \sum_{j=1}^{q_{i,k}} \phi_{i,k,j}^{(q_{i,k})} z_{t-j} \qquad (1)$$

for $\tau_{i,k} \le t < \tau_{i+1,k}$ and $i = 0,\dots,k$. Here, $\tau_{i,k}$ is the i^{th} change-point (defined as the index of the observation just before the i^{th} change-point, with the usual convention $\tau_{0,k} = 1$ and $\tau_{k+1,k} = n$, and for each i^{th} segment:

- $q_{i,k}$ and $\phi_{i,k}^{(q_{i,k})} = (\phi_{i,k,1}^{(q_{i,k})},\dots,\phi_{i,k,q_{i,k}}^{(q_{i,k})})'$ are the model order and the parameter of the MA model associated to this segment.
- z_t is the Gaussian noise of variance $\sigma_{i,k}^2$ associated to the MA model in this segment, i.e. $z_t \sim N(0,\sigma_{i,k}^2)$ for $\tau_{i,k} \le t < \tau_{i+1,k}$.

As shown in [1], for each the i^{th} segment ($i = 0,\dots, k$), the MA model of the order $q_{i,k}$ is invertible if $\phi_{i,k}^{(q_{i,k})}$ belongs to

$$I_{q_{i,k}} = \{\phi_{i,k}^{(q_{i,k})} \in \Re^{q_{i,k}} \Big| 1 + \phi_{i,k,1}^{(q_{i,k})} x + \dots$$

$$\ldots + \phi_{i,k,q_{i,k}}^{(q_{i,k})} x^{q_{i,k}} \neq 0, \quad x \in C, \quad |x| \leq 1\}.$$

The invertible region $I_{q_{i,k}}$ is difficult to be find if $q_{i,k} > 2$. To solve this problem, a reparameterization of a MA($q_{i,k}$) model is adopted [16]. The notation refers to the MA($q_{i,k}$) model of order $q_{i,k}$. For a MA($q_{i,k}$) model, There is a one-to-one transformation

$$G : \rho_{i,k}^{(q_{i,k})} \in]-1,1[^{q_{i,k}} \rightarrow \phi_{i,k}^{(q_{i,k})} \in I_{q_{i,k}}$$

where $\rho_{i,k}^{(q_{i,k})} = (\rho_{i,k,1}^{(q_{i,k})}, \ldots, \rho_{i,k,q_{i,k}}^{(q_{i,k})})'$ is the vector of the first $q_{i,k}$ inverted partial autocorrelations of the MA($q_{i,k}$) model (see [16] for a definition of the vector $\rho_{i,k}^{(q_{i,k})}$). The results of this reparametrization, show the condition that $\phi_{i,k}^{(q_{i,k})} \in I_{q_{i,k}}$ becomes $|\rho_{i,k}^{(q_{i,k})}| < 1$ (j = 1,…,$q_{i,k}$).

Let $\theta = (k, \tau^{(k)}, q^{(k)}, \phi^{(k)}, \sigma^{(k)})'$ be a parameter vector. Here $\tau^{(k)} = (\tau_1, \ldots, \tau_k)'$, $q^{(k)} = (q_{0,k}, \ldots, q_{k,k})'$, $\phi^{(k)} = (\phi_{0,k}^{(q_{0,k})}, \ldots, \phi_{k,k}^{(q_{k,k})})'$ and $\sigma^{(k)} = (\sigma_{0,k}, \ldots, \sigma_{k,k})'$. Suppose that $y_t \, (t = 1, \ldots, n)$ is a random sample drawn from a population having a piecewise MA model. This paper proposes a reversible jump MCMC algorithm to estimate the parameter θ.

2. Research method

Here, the parameter θ is estimated by using Bayesian method. Unfortunately, the Bayesian estimator cannot be determined analytically because the likelihood function of the parameter θ has a complicated form. To overcome these problems, a reversible jump MCMC Algorithm [13] is used.

The parameter estimation is done in a bayesian framework. First, the function of probability is determined. Second, the prior distribution for the parameters of the piecewise model is selected. Third, the posterior distribution is determined by using Bayes's theorem. Fourth, since the bayes estimator for parameters can not be calculated analytically, the bayes estimator for the parameters is calculated using the MCMC reversible jump algorithm. The MCMC reversible jump algorithm is implemented in three stages: (a) The birth of the change-point number, the death of the change-point number, and (c) the change of location of the change-point. Furthermore, for the number of change-point and change-point location known there are three sub-stages, namely: birth order MA model, death order MA model, and changes in the coefficient value of the MA model.

3. Results and anaysis

3.1. Maximum Likelihood Function

In lieu of the exact likelihood, an approximation to the likelihood is developed. Let q_{max} be the maximum number of order. Since the residual sequence is a normal distribution, the approximate likelihood takes a form :

$$f(s|\theta) = \prod_{i=0}^{k} (2\pi\sigma_{i,k}^2)^{-\frac{1}{2}(\tau_{i+1,k} - \tau_{i,k})} \qquad (2)$$

$$\exp - \frac{1}{2\sigma_{i,k}^2} \sum_{t=\tau_{i,k}+1}^{\tau_{i+1,k}} (y_t - \sum_{j=1}^{q_{i,k}} G(\rho_{i,k}^{(q_{i,k})}) \hat{z}_{t-j})^2$$

where

$$s = (y_{q_{max}+1}, \ldots, y_n)'$$

and with letting

$$\hat{z}_1 = \ldots = \hat{z}_{q_{max}} = 0$$

the t^{th} residual ($t = q_{max}+1, \ldots, n$) is calculated by

$$\hat{z}_t = y_t - \sum_{j=1}^{q_{i,k}} G(\rho_{i,k}^{(q_{i,k})}) \hat{z}_{t-j}, \qquad \tau_{i,k} \leq t < \tau_{i+1,k}, \quad i = 0, \ldots, k.$$

3.2. Bayesian Approach

A Bayesian approach is adopted in this work. It implies the choice of priors. Denote k_{max} as the maximum number of change-point. The number k of positions is drawn following a binomial distribution $B(k_{max}, \lambda)$ with parameter $\lambda (0 \leq \lambda \leq 1)$ and k_{max}.

For k fixed, the change-point positions are distributed as the even numbered positions of the order statistics of 2k+1 points uniformly drawn without repetitions in {2,…,n}. This choice avoids too small interval between changes. We obtain (k<<n large). The model order $q_{i,k}$ (i=0,…,k) is independent with the same binomial distribution $B(q_{max}, \mu)$ with parameter $\mu (0 \leq \mu \leq 1)$ and q_{max}. For $q_{i,k}$ fixed, where $\rho_{i,k,j}^{(q_{i,k})}$ is independent with the same uniform distribution $\rho_{i,k,j}^{(q_{i,k})} \sim U(-1,1)$ and $\sigma_{i,k}^2$ is independent and distributed according to $IG(\alpha/2, \beta/2) \, (\alpha > 0, \beta > 0)$. In order to have the robust prior, we consider the hyper-parameter vector $(\lambda, \mu, \alpha, \beta)$ to be random. The hyper-parameters λ and ν are drawn following the same uniform prior on (0,1), i.e. $\lambda \sim U(0,1)$ and $\mu \sim U(0,1)$. Set $\alpha = 2$ and we choose a non-informative improper Jeffreys prior for β. Then the prior distribution of the parameters θ is given by:

$$\pi(\theta;\xi) = \pi(k;\lambda)\pi(\tau^{(k)};k)\pi(q^{(k)};k)\pi(\rho^{(k)} \\ k, q^{(k)})\pi(\sigma^{2(k)};\alpha,\beta,k) \qquad (3)$$

where $\xi = (\lambda, \mu, \beta)$. By the classical Bayesian formula, the posterior distribution is

$$\pi(\theta, \xi|s) \propto f(s|\theta)\pi(\theta;\xi) \qquad (4)$$

That is the product of likelihood function in (2) and prior distribution in (3). The Bayesian inference for parameter (θ, ξ) is based on the posterior distribution. This posterior distribution is calculated by (4). In our case, it is not analytically possible to obtain this quantity. The reversible jump MCMC algorithm is therefore applied.

3.3. Reversible Jump MCMC

The key idea is to build an ergodic Markov chain $(\theta, \xi)_{(j)}$ $(j = 1, \ldots, M)$ whose equilibrium distribution is the posterior distribution $\pi(\theta, \xi|s)$. This sample generated by the Markov chain can be used to estimate all posterior features of interest. If the number of change-points k and the model orders $q^{(k)}$ are known, then the Metropolis-Hasting algorithm can be used to simulate a process according to this posterior distribution. In our case, since both k and $q^{(k)}$ are unknown, the chain must jump from the model $(k, q^{(k)})$ with parameters $(\tau^{(k)}, \rho^{(k)}, \sigma^{2(k)})$ to the model $(k', q^{(k')})$ with parameters $(\tau^{(k')}, \rho^{(k')}, \sigma^{2(k')})$. Green [13] has proposed a solution to such problems of model selection. This model selection is done in two stages: First stage, the reversible jump MCMC algorithm is used to define the jump between models of differing dimensionality in term of k. In this work, the moves are chosen to be: birth of a change-point, death of a change-point and update of the change-point positions. Second stage, for k fixed we use the reversible jump MCMC algorithm to define the jump between models of differing dimensionality, but in each term of $q_{i,k}$ (i = 1,…,k). For the moves, the following transitions are used: birth of the model parameter, death of the model parameter and update of the parameter.

Using the Markov chain previously defined, to simulate (after a burn-in period) a random vector distributed as the posterior distribution $\pi(\theta, \xi|s)$. In the proposed implementation, the samples $k_{(j)}$ from the joint posterior distribution $\pi(\theta, \xi|s)$ are collected after ignoring the other parameters. This strategy provides the marginal distribution of the change-point number k. Consequently, the marginal maximum posteriori estimator of parameter k can be easily determined:

$$\hat{k} = \arg\max \hat{p}[k_{(j)} = l|s] \quad k_{(j)} \in \{0,\ldots,k_{max}\}$$

Once parameter has been estimated, the change-point locations and the orders can be estimated as follows:

$$\tau^{(k)} = \arg\max \hat{p}[\tau_{(j)}^{(k)} | k = \hat{k}, s] \quad \tau_{(j)}^{(k)} \in \{2,\ldots,n-1\}_k$$

and

$$\hat{q}_{i,k} = \arg\max \hat{p}[q_{i,k_{(j)}} | k = \hat{k}, s] \quad q_{i,k_{(j)}} \in \{0,\ldots,q_{max}\}$$

For $k = \hat{k}$ and $q_{i,k} = \hat{q}_{i,k}$ $(i = 1,\ldots,\hat{k})$, the MA parameter and the noise variance associated are estimated by the same way (using marginal maximum posteriori).

3.4. Simulation

The simulation results are presented for a synthetic signal. 250 samples of synthetic signal are generated from a signal model in (1) whose parameters are k = 1, $\tau^{(1)} = 125$, and the model order, the MA parameter and the noise variance for each segment are summarized in the Table 1.

Table 1: The parameter of the synthetic signal

i^{th} segment	$\sigma_{i,1}$	$q_{i,1}$	$\phi_{i,1}^{(q_{i,1})}$
0	0.5	1	0.7826
1	1.5	3	(0.52, -0.08, -0.96)

The MCMC simulation is run for 60.000 iterations, after a burn-in period of 10.000 iterations $(k_{max} = 10, q_{max} = 15)$. The histogram of the marginal a posteriori distribution k is plotted in the Fig. 1, and we obtain the marginal maximum posteriori of k. Here $\hat{k} = 1$.

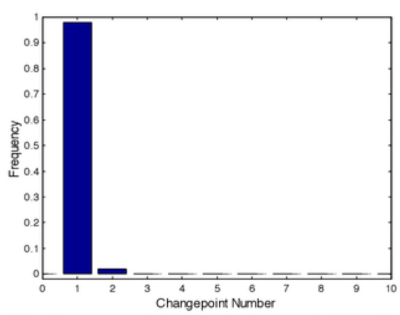

Fig. 1: The histogram of the marginal posteriori distribution of k

The Marginal Maximum Posteriori of $\tau^{(1)}$ is obtained. Here $\hat{\tau}^{(1)} = 127$. This estimated change-point and the syntetic signal data are plotted in the Fig. 2.

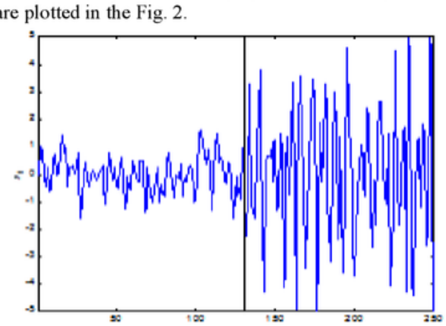

Fig. 2: The estimated change-point and the synthehtic signal

From the Fig. 2, it can see that this signal data set is composed of two different MA models.

Then for $\hat{k} = k$ fixed, the histogram of the conditional marginal posterior distribution of $\tau^{(1)}$, $q_{0,1}$ and $q_{1,1}$ is given in the Fig. 3, Fig. 4 and Fig. 5 respectively.

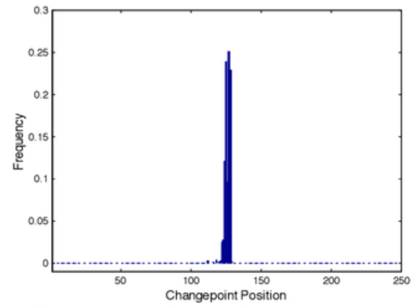

Fig. 3: The histogram of the conditional marginal posteriori distribution of $\tau^{(1)}$.

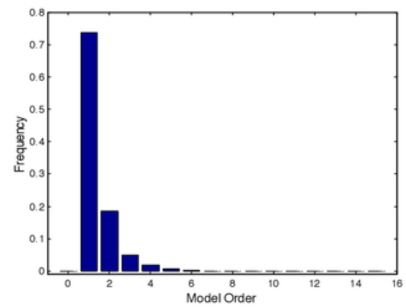

Fig. 4: The histogram of the conditional marginal posteriori distribution of $q_{0,1}$.

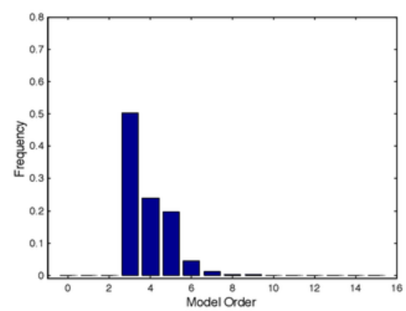

Fig. 5: The histogram of the conditional marginal posteriori distribution of $q_{0,1}$.

From the Fig. 4 and Fig. 5, we obtain the maximum marginal posterior of $q_{0,1}$ and $q_{1,1}$ are $\hat{q}_{0,1} = 1$ and $\hat{q}_{1,1} = 3$. Given $q_{0,1} = \hat{q}_{0,1}$, the Curve a (Fig. 6) shows the conditional marginal posterior distribution of $\phi_{0,1,1}^{(q_{0,1})}$ by using Gaussian kernel with the standard deviation 0.2. Similarly for $q_{1,1} = \hat{q}_{1,1}$, in the same Fig. 6 (Curve b, Curve c and Curve d) shows the conditional marginal posterior distribution of $\phi_{1,1,1}^{(q_{1,1})}$, $\phi_{1,1,2}^{(q_{1,1})}$ and $\phi_{1,1,3}^{(q_{1,1})}$ by using the same Gaussian kernel with the standard deviation 0.2.

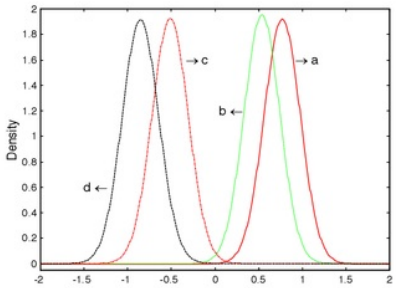

Fig. 6: The histogram of the conditional marginal posteriori distribution of the coefficient of the MA models.

Finally, the histogram of the marginal posteriori distribution of $\sigma_{0,1}$ and $\sigma_{1,1}$ are given in the Fig. 7 and Fig. 8.

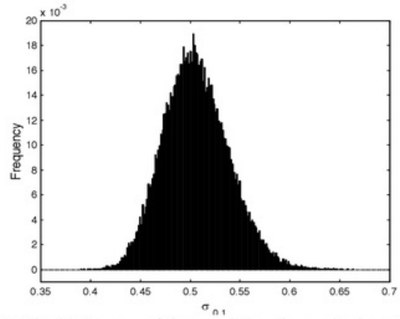

Fig. 7: The histogram of the conditional marginal posteriori distribution of $\sigma_{0,1}$.

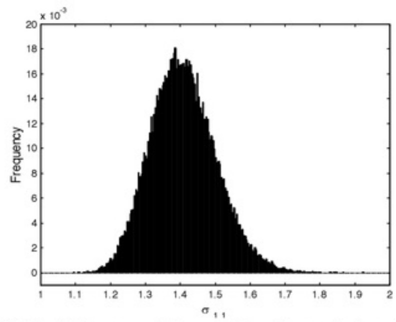

Fig. 8: The histogram of the conditional marginal posteriori distribution of $\sigma_{1,1}$.

4. Conclusion

This paper studied a change-point detection of the MA model based on reversible jump MCMC algorithm. The first algorithm generated Markov chain samples distributed according to the joint posterior distribution of the unknown parameters. These samples were then used to derive the maximum marginal posterior estimators. This algorithm showed good performance for the change-point detection of the MA model. A comparison with other existing approaches on real signal data is currently under investigation.

Acknowledgement

The author would like to thank Dr. Michel Doisy, University of Paul Sabatier (France), for his suggestions to improve this paper.

References

[1] H. Nakamura, Y. Toquenaga, "Estimating Colony Locations of Bumble Bees with Moving Average Model", Ecological Research, Vol. 17, Issue 1, pp. 39-48, 2002.
[2] Y. Wang, F. Ding, "Iterative Estimation for a Non-Linear IIR Filter with Moving Average Noise by Means of the Data Filtering Technique", IMA Journal of Mathematical Control & Information, Vol. 34, Issue 3, pp. 745-764, 2017.
[3] C. Dimitriou-Fakalou, "Yule-Walker Estimation for the Moving Average Model", International Journal of Stochastic analysis, pp. 1-20, 2011.
[4] A. Hong, A. Chen, "Piecewise Regression Model Construction with Sample Efficeint Regression Treen (SERT) and Applications to Semiconductor Yield Analysis", Journal of Process Control, Vol. 22, Issue 7, pp. 1307-1317, 2012.
[5] T.M.O. Diallo, A.J.S. Morin, "Power of Latent Growth Curve Models to Detect Piecewise Linear Trajectories", Structural Equation Modeling, Vol. 22, Issue. 3, pp. 449-460, 2015.
[6] B. Tang, Y. Xiao, J. Wu, "A Piecewise model of Virus-Immune System with Two Thresholds", Mathematical Biosciences, Vol. 278, pp. 63-76, 2016.
[7] B. Tang, Y. Xiao, S. Sivalganathan, J. Wu, "A Piecewise Model of Virus-Immune System with Effector Cell-Guided Therapy", Applied Mathematical Modelling, Vol. 47, pp. 227-248, 2017.
[8] F. Chamroukhi, "Piecewise Regression Mixture for Simultaneous Functional Data Clustering and Optimal Segmentation", Journal of Classification, Vol. 33, Issue 3, pp. 374-411, 2016.
[9] Z-H. Xu, X-A. Yin, C. Zhang, Z.F. Yang, "Piecewise Model for Species-discharge Relationships in Rivers", Ecological Engineering, Vol. 96, pp. 208-213, 2016.
[10] M-J. Buscot, S.S. Wotherpoon, C.G. Magnussen, M. Juonala, M.A. Sabin, D.P. Burger, T. Lehtimaki, J.S.A. Viikari, N. Hutri-Kahonen, O.T. Raitakari, R.J. Thomson, "Bayesian Hierarchical Piecewise Regression Models : A Tool to detect Trajectory Divergence Between Groups in Long-Term Observational Studies", BMC Medical Research Methodology, Vol. 17, pp. 1-15, 2017.
[11] M.R-U. Hector, M. Garnotel, M. Doron, P. Jallon, G. Charpentier, S. Franc, E. Huneker, C. Simon, S. Bonnet, "An Original Piecewise Model for Computing Energy Expenditure from Accelerometer and Heart Rate Signals", Physiological Measurement, Vol. 38, Issue 8, pp. 1-1, 2017.
[12] I. Soleimanmeigouni, X. Xiao, A. Ahmadi, M. Xie, A. Nissen, U. Kumar, "Modeling the Evolution of Ballasted Railway Track Geometry by Two-Level Piecewise Model", Structure & Infrastructure Engineering : Maintenance, Management, Life-Cycle Design & Performance, Vol. 14, Issue 1, pp. 33-45, 2018.
[13] P.J. Green, "Reversible Jump MCMC Computation and Bayesian Model Determination", Biometrika, Vol. 82, pp. 711-732, 1995.
[14] Suparman, M. Doisy, "Hierarchical Bayesian of ARMA Models Using Simulated Algorithm", Journal of Telkomnika, Vol. 12, No. 1, pp. 87-96, 2014.
[15] Suparman, "New Segmentation of Piecewise Linear Regression Models Using Reversible Jump MCMC Algorithm", Journal of Computer Technology and Application, Vol. 6, No. 1, pp. 14-18, 2015.
[16] R.J. Bhansali, "The inverse partial correlation function of a time series and its applications", J.Multivar. Anal., vol. 13, pp. 310-327, 1983.

Change-Point Detection in Moving Average Model Using Reversible Jump MCMC Algorithm

ORIGINALITY REPORT

15%

SIMILARITY INDEX

PRIMARY SOURCES

1	www-sigproc.eng.cam.ac.uk Internet	112 words — 4%
2	S. Suparman, Michel Doisy, Jean- Yves Tourneret. "Changepoint detection using reversible jump MCMC methods", IEEE International Conference on Acoustics Speech and Signal Processing, 2002 Crossref	93 words — 4%
3	Mohammad Rafighi, Abdulkadir Güllü. "Design of a novel walking assistance device for people with walking disabilities", International Journal of Engineering & Technology, 2017 Crossref	38 words — 1%
4	www.ciic.cms.net.my Internet	30 words — 1%
5	linux.web.cern.ch Internet	18 words — 1%
6	www.mattababy.org Internet	14 words — 1%
7	K. Ko. "Bayesian Wavelet-Based Methods for the Detection of Multiple Changes of the Long Memory Parameter", IEEE Transactions on Signal Processing, 11/2006 Crossref	12 words — < 1%
8	Vikram Saini, Lillie Dewan. "Instrument variable method based on	

nonlinear transformed instruments for Hammerstein system identification", Journal of Vibration and Control, 2017
Crossref
10 words — < 1%

9 www.neas-seminars.com
Internet
9 words — < 1%

10 W. Ng. "Wideband Array Signal Processing Using MCMC Methods", IEEE Transactions on Signal Processing, 2/2005
Crossref
9 words — < 1%

11 eprints.soton.ac.uk
Internet
8 words — < 1%

12 Zhou, X.. "Construction of genomic networks using mutual-information clustering and reversible-jump Markov-chain-Monte-Carlo predictor design", Signal Processing, 200304
Crossref
8 words — < 1%

13 B. Roe, I. Sheldon. "Credence Good Labeling: The Efficiency and Distributional Implications of Several Policy Approaches", American Journal of Agricultural Economics, 2007
Crossref
7 words — < 1%

14 M. Sambridge, K. Gallagher, A. Jackson, P. Rickwood. "Trans-dimensional inverse problems, model comparison and the evidence", Geophysical Journal International, 2006
Crossref
7 words — < 1%

15 Amr Hussein, Hossam Kasem, Mohamed Adel. "Efficient spectrum sensing technique based on energy detector, compressive sensing, and de-noising techniques", International Journal of Engineering & Technology, 2016
Crossref
6 words — < 1%

16 Christophe Andrieu. "Robust Full Bayesian Learning for Radial Basis Networks", Neural Computation,
6 words — < 1%

10/2001

Crossref

EXCLUDE QUOTES	OFF	EXCLUDE MATCHES	OFF
EXCLUDE BIBLIOGRAPHY	ON		